

# (12) UK Patent Application (19) GB (11) 2 285 445 (13) A

(43) Date of A Publication 12.07.1995

(21) Application No 9324955.5	(51) INT CL <sup>6</sup> C12N 15/11 , C12Q 1/68
(22) Date of Filing 06.12.1993	(52) UK CL (Edition N ) C3H HB5 G1B BAH B620
(71) Applicant(s) PNA Diagnostics A/S  (Incorporated in Denmark)  Lerso Park Alle 42, DK-2100 Copenhagen O, Denmark	(56) Documents Cited WO 92/20703 A1 Health Periodicals Abstract No.11677016,& Science, Vol.254, (Dec.6 1991),pages 1497-1500.
(72) Inventor(s) Christopher John Stanley Henrik Orum Mikkel Jorgensen Ole Basbøll	(58) Field of Search UK CL (Edition M ) C3H HB4B HB5 HB7X , G1B BAH BAX INT CL <sup>5</sup> C12N 15/11 , C12Q 1/68 ONLINE DATABASES: WPI, DIALOG (BIOTECH)
(74) Agent and/or Address for Service W H Beck, Greener & Co 7 Stone Buildings, Lincoln's Inn, LONDON, WC2A 3SZ, United Kingdom	

## (54) Protecting nucleic acids and methods of analysis

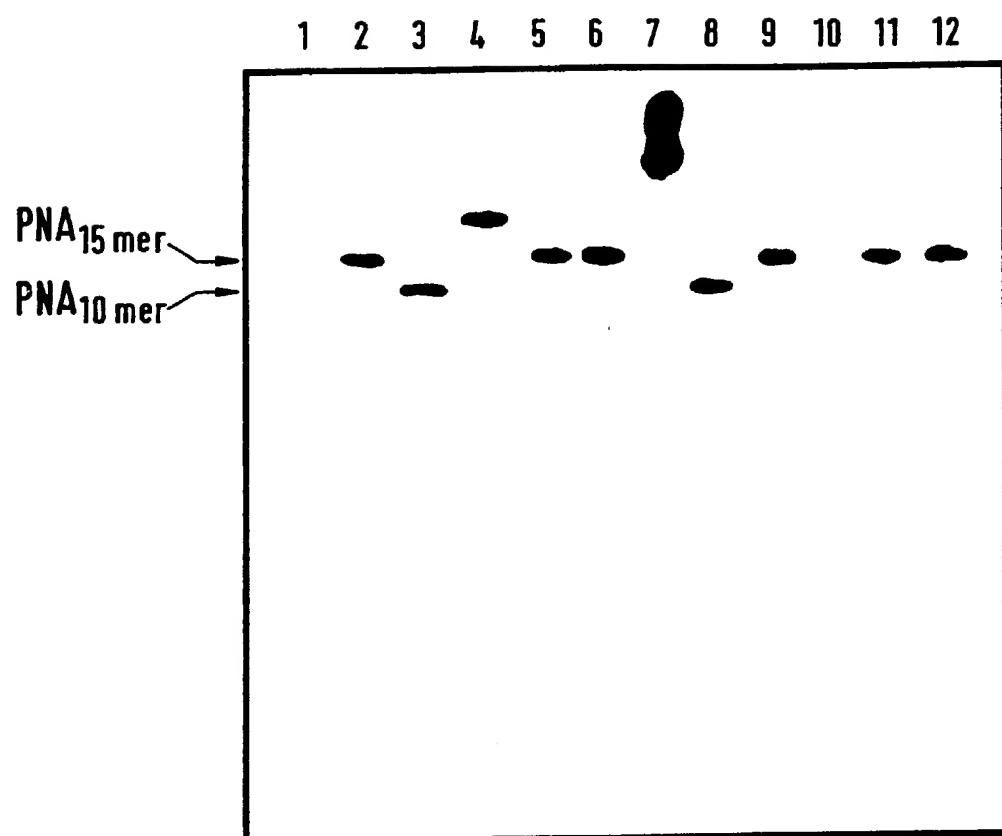
(57) A selected region of a nucleic acid is protected from attack by a reagent e.g. a nuclease by complexing to the nucleic acid a nucleic acid analogue which hybridises thereto in a sequence selective manner. The surviving sequence may be detected in an assay, optionally after amplification. A PCR reaction may be sterilised and its product assayed by protecting a characterising region of the product by hybridisation as above followed by nuclease degradation of unprotected nucleic acid. The nucleic acid analogue is most preferably a polymer e.g. a polyamide, polythioamide, polysulphinamide or polysulphonamide to the backbone of which is a bound a sequence of ligands.

GB 2 285 445 A

46 43 95

1 / 1

Fig. 1



PROTECTING NUCLEIC ACIDS AND METHODS OF ANALYSIS

The present invention relates to methods of protecting nucleic acids from attack, eg. by nucleases, and analytical 5 methods and in particular to methods for detecting the presence in a nucleic acid containing sample of a sequence of interest.

At present the usual way of determining whether a particular base sequence is present in a sample of nucleic 10 acid is filter hybridisation. The nucleic acid is extracted from starting materials and purified. It may need to be amplified by a process such as PCR. The nucleic acid is then immobilised in single stranded form on a filter and is probed with a labelled oligonucleotide having a sequence complementary 15 to that which is to be detected. Before being immobilised, the nucleic acid may be run on a gel to separate different molecular weight nucleic acid fragments. This is the widely used Southern blot procedure. This procedure has numerous drawbacks. The nucleic acid needs to be separated 20 from other materials. Where the nucleic acid is DNA, it will normally be necessary to go through a denaturation step. The stringency conditions of the hybridisation will need careful adjustment to avoid false positives and to achieve the desired level of discrimination between similar sequences. Because 25 of the limited signal to noise of the filter hybridisation, amplification steps will often be necessary to produce sufficient of the sequence of interest for detection. This generates many opportunities for the production of false negatives.

30 Nucleic acid analogues having important new utilities in assay procedures and in the field of diagnostics have been described in WO 92/20703. These nucleic acid analogues had a number of new properties making them of special importance in the field of diagnostics as well as in the field of anti-sense therapeutics. Such nucleic acid analogues are referred 35 to herein as "PNAs".

They typically feature a polyamide backbone bearing a sequence of ligands which are nucleic acid bases. The analogues described there have the property of hybridising with great specificity and stability to natural nucleic acids 5 of complementary sequence.

We have now discovered that such nucleic acid analogues have the further valuable property of protecting sequences of single stranded nucleic acids to which they are hybridised from being degraded by certain reagents under circumstances 10 and conditions in which the reagents would normally be effective to produce degradation of said single stranded nucleic acids. Regions of the nucleic acid not hybridised to the nucleic acid analogue may be degraded as normal.

Accordingly, the present invention provides a method of 15 protecting a selected region of a nucleic acid from attack by a reagent, which method comprises forming a complex between the nucleic acid and a nucleic acid analogue which hybridises thereto in a sequence selective manner and exposing said nucleic acid to a nucleic acid attacking reagent, wherein said 20 complex is more stable against attack by said reagent than said starting nucleic acid.

The nucleic acid may be RNA or DNA.

Preferably, said reagent is a nuclease and is capable of degrading said nucleic acid either completely or by cleaving 25 it at certain locations.

In the most straightforward case, the selected region of the nucleic acid to be protected will be constituted by the nucleic acid sequence to which the nucleic acid analogue hybridises. However, where the attacking reagent is an 30 exonuclease capable of degrading nucleic acids under the conditions used only by working in from the 5' and 3' ends, a larger selected region may be protected by placing a respective nucleic acid analogue at each end of the selected region leaving part of the selected region unhybridised 35 between the two nucleic acid analogue sequences. The intervening part of the selected region is fenced off from

access by the exonuclease. Where the exonuclease is, under the conditions employed, able only to attack from the 5' or 3' end, a single nucleic acid analogue sequence may be employed to fence off all that part of the nucleic acid which 5 lies 3' or 5' respectively to the nucleic acid analogue.

Where an endonuclease is employed as the attacking reagent, a short stretch of nucleic acid exposed unhybridised between two nucleic acid analogue oligomers lying adjacent one another on the nucleic acid may none the less form part of the 10 selected region which is protected, by virtue of some interaction between the end bases of the two nucleic acid analogue sequences or because of interference by the nucleic acid analogue with the activity of the endonuclease.

The nucleic acid analogue is preferably one comprising 15 a polymeric strand which includes a sequence of ligands bound to a backbone made up of linked backbone moieties, which analogue is capable of hybridisation to a nucleic acid of complementary sequence.

Said nucleic acid analogue backbone is preferably a 20 polyamide, polythioamide, polysulphinamide or polysulphonamide backbone.

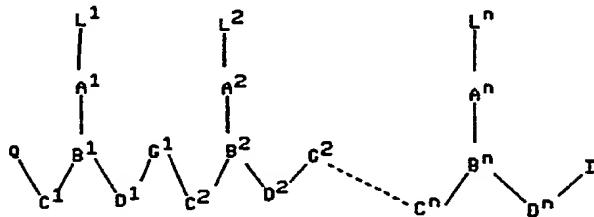
Preferably, said linked backbone moieties are peptide bonded amino acid moieties.

The nucleic acid analogue is preferably capable of 25 hybridising to a nucleic acid of complementary sequence to form a hybrid which is more stable against denaturation by heat than a hybrid between the conventional deoxyribonucleotide corresponding in sequence to said analogue and said nucleic acid.

30 Preferably, said nucleic acid analogue is a peptide nucleic acid in which said backbone is a polyamide backbone, each said ligand being bonded directly or indirectly to an aza nitrogen atom in said backbone, and said ligand bearing nitrogen atoms mainly being separated from one another in said 35 backbone by from 4 to 8 intervening atoms.

Preferably also, nucleic acid analogue is capable of hybridising to a double stranded nucleic acid in which one strand has a sequence complementary to said analogue, in such a way as to displace the other strand from said one strand.

5 Preferably, the nucleic acid analogue has the general formula 1:



Formula 1

wherein:

n is at least 2,  
10 each of L<sup>1</sup>-L<sup>n</sup> is independently selected from the group consisting of hydrogen, hydroxy, (C<sub>1</sub>-C<sub>4</sub>)alkanoyl, naturally occurring nucleobases, non-naturally occurring nucleobases, aromatic moieties, DNA intercalators, nucleobase-binding groups, heterocyclic moieties, and reporter ligands, but  
15 normally at least one L will be a nucleobase binding group such as a naturally occurring nucleobase and preferably at least 90% of the groups L will be such nucleobase binding groups;

each of C<sup>1</sup>-C<sup>n</sup> is (CR<sup>6</sup>R<sup>7</sup>)<sub>y</sub> where R<sup>6</sup> is hydrogen and R<sup>7</sup> is  
20 selected from the group consisting of the side chains of naturally occurring alpha amino acids, or R<sup>6</sup> and R<sup>7</sup> are independently selected from the group consisting of hydrogen, (C<sub>2</sub>-C<sub>6</sub>)alkyl, aryl, aralkyl, heteroaryl, hydroxy, (C<sub>1</sub>-C<sub>6</sub>)alkoxy, (C<sub>1</sub>-C<sub>6</sub>)alkylthio, NR<sup>3</sup>R<sup>4</sup> and SR<sup>5</sup>, where R<sup>3</sup> and R<sup>4</sup>  
25 are as defined below, and R<sup>5</sup> is hydrogen, (C<sub>1</sub>-C<sub>6</sub>)alkyl, hydroxy, alkoxy, or alkylthio-substituted (C<sub>1</sub> to C<sub>6</sub>)alkyl or

R<sup>6</sup> and R<sup>7</sup> taken together complete an alicyclic or heterocyclic system;

each of D<sup>1</sup>-D<sup>n</sup> is (CR<sup>6</sup>R<sup>7</sup>)<sub>z</sub> where R<sup>6</sup> and R<sup>7</sup> are as defined above;

5 each of y and z is zero or an integer from 1 to 10, the sum y + z being from 2 to 10 (preferably being more than 2, and most preferably each of y and z being 1 or 2;

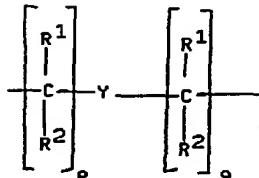
each of G<sup>1</sup>-G<sup>n-1</sup> is -NR<sup>3</sup>CO-, -NR<sup>3</sup>C<sup>S</sup>-, -NR<sup>3</sup>SO- or -NR<sup>3</sup>SO<sub>2</sub>-, in either orientation, where R<sup>3</sup> is as defined below;

10 each of A<sup>1</sup>-A<sup>n</sup> and B<sup>1</sup>-B<sup>n</sup> are selected such that:

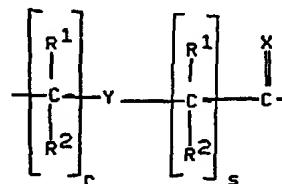
(a) A is a group of formula (IIa), (IIb), (IIc) or (IId), and B is N or R<sup>3</sup>N<sup>+</sup>; or

(b) A is a group of formula (IId) and B is CH;

15



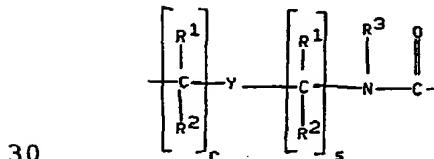
20



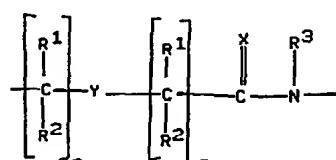
Formula IIa

Formula IIb

25



30



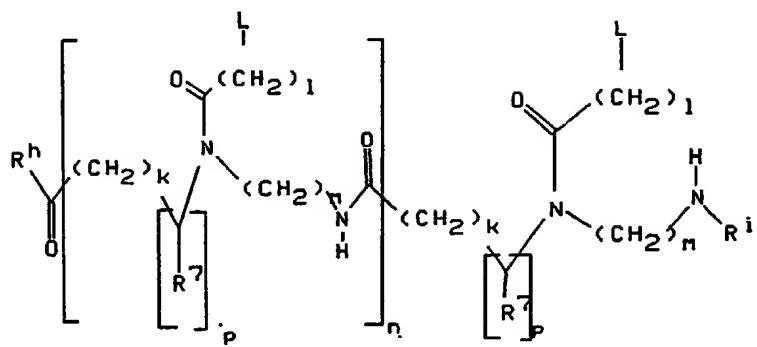
Formula IIc

Formula IId

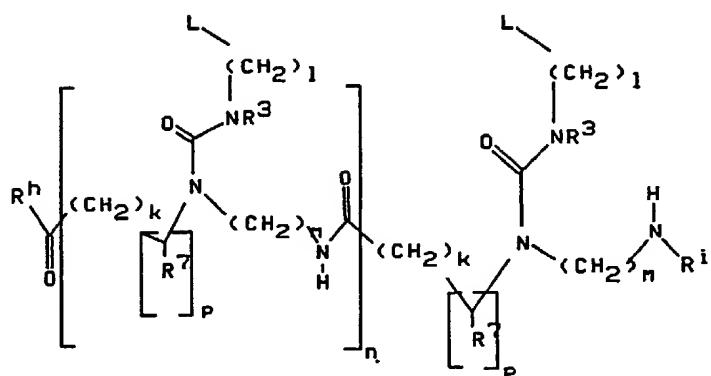
35

wherein:

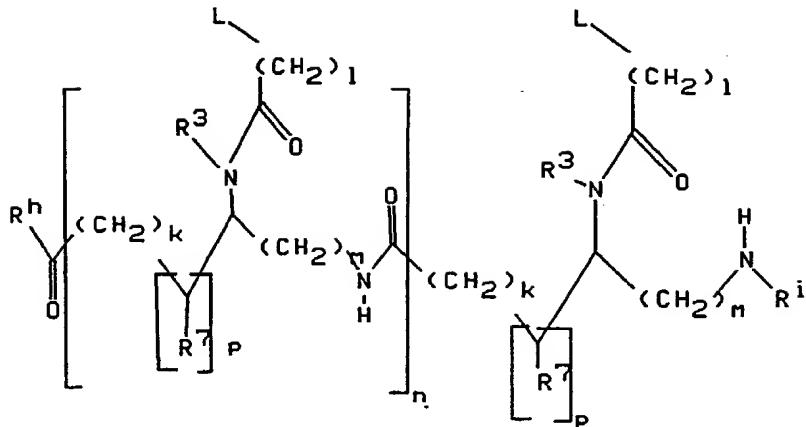
- X is O, S, Se, NR<sup>3</sup>, CH<sub>2</sub> or C(CH<sub>3</sub>)<sub>2</sub>;
  - Y is a single bond, O, S or NR<sup>4</sup>;
  - each of p and q is zero or an integer from 1 to 5, the sum p+q being not more than 10;
  - 5 each of r and s is zero or an integer from 1 to 5, the sum r+s being not more than 10;
  - each R<sup>1</sup> and R<sup>2</sup> is independently selected from the group consisting of hydrogen, (C<sub>1</sub>-C<sub>4</sub>)alkyl which may be hydroxy- or alkoxylthio-substituted, hydroxy, alkoxy, alkylthio, amino and halogen; and
  - 10 each R<sup>3</sup> and R<sup>4</sup> is independently selected from the group consisting of hydrogen, (C<sub>1</sub>-C<sub>4</sub>)alkyl, hydroxy- or alkoxy- or alkylthio-substituted (C<sub>1</sub>-C<sub>4</sub>)alkyl, hydroxy, alkoxy, alkylthio and amino;
  - 15 Q is -CO<sub>2</sub>H, -CONR'R'', -SO<sub>3</sub>H or -SO<sub>2</sub>NR'R'' or an activated derivative of -CO<sub>2</sub>H or -SO<sub>3</sub>H; and
  - I is -NR'R'', where R' and R'' are independently selected from the group consisting of hydrogen, alkyl, amino protecting groups, reporter ligands, intercalators, chelators, peptides, 20 proteins, carbohydrates, lipids, steroids, nucleosides, nucleotides, nucleotide diphosphates, nucleotide triphosphates, oligonucleotides, including both oligoribonucleotides and oligodeoxyribonucleotides, oligonucleosides and soluble 25 and non-soluble polymers.
- More preferably, said nucleic acid analogue comprises a compound of the general formula III, IV or V:



Formula III



Formula IV



Formula V

wherein:

each L is independently selected from the group consisting of hydrogen, phenyl, heterocyclic moieties,  
5 naturally occurring nucleobases, and non-naturally occurring nucleobases;

each R<sup>7</sup> is independently selected from the group consisting of hydrogen and the side chains of naturally occurring alpha amino acids;

10 n is an integer greater than 1,

each k, l, and m is, independently, zero or an integer from 1 to 5;

each p is zero or 1;

R<sup>h</sup> is OH, NH<sub>2</sub> or -NHLysNH<sub>2</sub>; and

15 R<sup>i</sup> is H or -COCH<sub>3</sub>.

The method of protecting a nucleic acid described above can be used in a method of detecting a nucleic acid sequence. Thus the present invention includes in a second aspect a method of detecting the presence in a nucleic acid sample of 20 a sequence, which method comprises exposing the sample of nucleic acid to a nucleic acid analogue capable of hybridising in sequence selective manner to said sequence if present in

said nucleic acid sample under hybridising conditions so as to form a complex between said nucleic acid analogue and a region of said nucleic acid containing said sequence, exposing said nucleic acid to a reagent capable of degrading said 5 nucleic acid under conditions such that the said region of said nucleic acid forming said complex is protected from attack by said reagent whilst the remainder of said nucleic acid is degraded, and detecting the presence of said complex.

There will be many ways of detecting the presence of said 10 complex. These may be divided into those which detect the complex as such and those in which the complex is broken up into its constituents of nucleic acid and nucleic acid analogue and the presence of the nucleic acid or the nucleic acid analogue is detected. The existence of the complex can 15 be deduced from the survival of the nucleic acid.

In the first group of methods, reliance may be placed upon differences in properties of the nucleic acid analogue and the nucleic acid/nucleic acid hybrid. Thus preferably, the presence of said complex is detected by 20 comparing the movement of said complex in electrophoresis with the movement of said nucleic acid analogue under similar conditions.

Preferably, the nucleic acid analogue or the nucleic acid carries a detectable label, suitable labels for the nucleic 25 acid analogue include a fluorescent label, a radio-label, an enzyme label, biotin, a spin label, a chemiluminescent label, an antigen label or an antibody label. The most preferred of these are the use of a fluorophore or a radioisotope as a label but generally any method of labelling applicable to 30 peptides can be used.

In a preferred practice of the second aspect of the invention, the nucleic acid analogue is a labelled nucleic acid analogue produced by a method for labelling a nucleic acid analogue comprising providing a nucleic acid analogue 35 with a peptide motif capable of functioning as a substrate for an enzyme in a labelling reaction and carrying out a said

labelling reaction comprising reacting the peptide motif of the nucleic acid analogue under the influence of an enzyme with a source of said label.

Preferably, said label is a radio-label.

5 The source of said label is preferably radio-labelled ATP.

Said enzyme is preferably a protein kinase.

The peptide motif is preferably the kemptide motif.

10 Preferably, the labelling reaction is a phosphorylation at a serine residue of said peptide motif.

An alternative method of labelling a nucleic acid analogue involves providing, preferably at one end of the nucleic acid analogue, a chelating moiety capable of binding at least one metal ion by chelation and chelating a 15 radioisotope or fluorophore directly or indirectly thereto. For instance, a europium ion may be chelated by a suitable chelating moiety carried by the nucleic acid analogue to act as a fluorescent label.

An alternative detection method in this first group makes 20 use of the nucleic acid sequence which is protected to act as a means of linking two nucleic acid analogue sequences which would otherwise not associate. If each bears a different type of label, the association of the two labels can be detected. By way of example, two nucleic acid analogue oligomers may be 25 hybridised to respective ends of a target nucleic acid sequence forming part of a larger nucleic acid sequence. One nucleic acid analogue is labelled with a moiety suitable for linking to a solid phase e.g. a chelating peptide motif such as His<sub>5</sub>. The other is provided with a detectable label, e.g. 30 is radio-labelled. After digestion of unprotected nucleic acid, the two nucleic acid analogue sequences are left joined via hybridisation to the target nucleic acid protected sequence. The complex may be captured on a solid support bearing chelatable nickel ions via chelation by the first 35 nucleic acid analogue. Radio-labelled excess second nucleic acid analogue may be washed away and the radio activity of the

second labelled nucleic acid analogue may be detected on the solid support or after elution from it.

Other examples would include having one nucleic acid analogue bearing biotin as a moiety suitable for linking to a solid phase instead of the chelating peptide motif and the second nucleic acid analogue linked to a detectable label, e.g. an enzyme or a fluorophore. If the detectable label is an enzyme capable of catalysing the production of a detectable product, the assay is analogous to the well established ELISA principle in immunoassay.

Alternative means for separating hybridised from non-hybridised nucleic acid analogue include reverse phase or ion exchange chromatography, or gel filtration. An antibody selection procedure could be used employing an antibody with specificity for intact nucleic acid or with specificity for the nucleic acid/nucleic acid hybrid.

In the second group of detection methods, the complex between the nucleic acid and the nucleic acid analogue is denatured and the survival of the protected nucleic acid sequence is demonstrated, suitably by known nucleic acid sequence detection methods such as a hybridisation assay. Optionally, the nucleic acid sequence may be subjected to an amplification procedure. LCR will generally be preferred for amplifying short protected sequences and PCR for longer ones.

Nucleic acid amplification procedures, especially the well known PCR procedures often encounter difficulties when the target sequence to be amplified is present amongst a large amount of other nucleic acid.

As is shown above, the present invention provides a ready means of "cleaning up" a nucleic acid sample by degrading all nucleic acid present except the target sequence, thus allowing amplification to be more easily carried out whether for the purposes of assay or preparatively.

Accordingly, the present invention provides in a third aspect, a method of carrying out a nucleic acid amplification comprising protecting a selected region of a nucleic acid

within a nucleic acid sample by hybridising to said nucleic acid a nucleic acid analogue which hybridises thereto in a sequence specific manner, exposing said sample to a nucleic acid attacking reagent to degrade the nucleic acid in the 5 sample except for said selected region, and amplifying said selected region.

To free the protected sequence for amplification, the nucleic acid/nucleic acid analogue hybrid may be denatured by heat. Any amplification method may be employed, including all 10 presently known method such as PCR and its variants, LCR and 3SR.

In WO (PCT/EP93/01435) a method is described for sterilising a nucleic acid amplification by hybridising to the amplification product a nucleic acid analogue which 15 will prevent any amplification from acting as an amplifiable target. This may be still further improved according to a fourth aspect of the present invention by protecting a region of the amplification product which excludes at least one primer binding site necessary for repetition of the 20 amplification by a method in accordance with the first aspect of the invention, degrading the remaining nucleic acid, including the primer binding sites of the amplification product, and detecting the survival of the protected nucleic acid as discussed above. Preferably, at least some, more 25 preferably the whole, of all the primer binding sequences are degraded.

The nucleic acid analogue may be present throughout a PCR procedure, provided that the temperatures employed are high enough to melt the nucleic acid analogue off the nucleic acid, 30 or may be added at the end of the amplification. The attacking reagent will need to be added following the amplification but may be added from a separate compartment within a closed amplification apparatus so that amplified product never has an opportunity to escape and become a 35 contaminant.

The invention will be illustrated by the following examples with reference to the accompanying drawings in which:

Figure 1 is an autoradiograph of a gel showing the results of Examples 3 to 5.

5 One convenient method of producing a PNA bearing a radio-label of high specific activity particularly preferred for use in the present invention is described in our co-pending British patent application bearing agent's reference P3685GB, filed on the same day as this application. According to such  
10 a method in its preferred form, a PNA is prepared having a terminal kemptide extension which undergoes a reaction with  $^{32}\text{P}$  ATP mediated by protein kinase A to acquire a  $^{32}\text{P}$  label bound to a serine residue. PNA's having peptide extensions of desired amino acid sequences can conveniently be produced  
15 by the Boc or Fmoc solid phase techniques well understood in the art once a starting PNA sequence has been built-up on the solid support suitably by using the Boc solid phase synthesis described in WO92/20703.

According to preferred forms of the first aspect of the  
20 present invention, a nucleic acid containing sequences to be protected may be treated with a PNA having a sequence complementary to the sequence of interest. If the nucleic acid is DNA, it may be in double stranded form and one may rely on the ability of the nucleic acid analogue to hybridise  
25 to its target sequence in double stranded form. However, it may be necessary and will normally be desirable to denature the double stranded target DNA, usually by heating. The DNA may be quite impure, i.e. other cell products may be present. Also, the DNA may be whole genomic DNA or total cell RNA  
30 extracted from cells and not digested into shorter lengths. The DNA or RNA could also be the product of a nucleic acid amplification procedure such as LCR or PCR or a NASBA (3SR) amplification. Preferred PNA's will bind sequence selectively with high affinity, even if the target DNA is double stranded.  
35 The nucleic acid sample may then be treated with a nuclease or nuclease mixture to degrade all of the nucleic acid not

protected by hybridisation to the PNA. The PNA may be present in substantial excess to ensure that if there is even one correct sequence present in the nucleic acid sample, it will be hybridised and protected.

5       The reaction product may be subjected to electro-phoresis on a gel or in a capillary to separate PNA which is hybridised to nucleic acid from non-hybridised PNA. Because PNA is an essentially neutral molecule, having one positive charge per molecule typically at a terminal amine group, it shows little  
10      tendency to migrate during electrophoresis. Nucleic acids on the other hand are highly negatively charged and hybridisation of PNA to a nucleic acid fragment results in a negatively charged hybrid which will migrate under electrophoresis in the opposite direction to the PNA itself. This effects a ready  
15      and highly specific separation of the hybridised and non-hybridised PNA. If the PNA is appropriately labelled, it is possible to detect extremely small quantities of it and the analytical method according to the second aspect of the invention provides an excellent signal to noise ratio.

20      PNA's are able to distinguish nucleic acid sequences differing by only one residue. Accordingly, the methods described above offer the possibility of a very quick and simple assay capable of determining whether a sample of genomic DNA contains a particular allelic variation within a  
25      particular gene. Furthermore, as PNA's of different lengths hybridised to nucleic acids are readily separable under electrophoresis, the possibility also exists of probing a nucleic acid sample with a plurality of PNA's simultaneously, digesting away all the unhybridised nucleic acid and by  
30      electrophoresis separating all the PNA/nucleic acid hybrids from the unhybridised excess PNA'S and from each other. One may thus in a single assay demonstrate that a DNA sample contains, possibly on different genes, some particular combination of allelic variations of interest, eg. one found  
35      to give rise to a particular phenotype such as a disease state.

Example 1

Preparation of a PNA-kemptide chimera.

5       The solid phase PNA synthesis described in  
WO 92/20703 was used to build up the sequence:  
Boc-NH(CH<sub>2</sub>)<sub>5</sub>CONH-TG(Z)T.A(Z)C(Z)G(Z).TC(Z)A(Z).C(Z)A(Z)  
A(Z).C(Z)TA(Z)-CONH-resin

10      The N terminal Boc group was removed by treatment with  
TFA and used as a starting point for a standard boc type solid  
phase peptide synthesis of the kemptide motif via the linker  
6-amino-hexanoic acid to produce the chimera:-

15      Boc-Leu-Arg(Tos)-Arg(Tos)-Ala-Ser-(Bzl)-Leu-Gly-NH(CH<sub>2</sub>)<sub>5</sub>CONH-  
TG(Z)T.A(Z)C(Z)G(Z).TC(Z)A(Z).C(Z)A(Z) A(Z).C(Z)TA(Z)-CONH-  
resin

20      The protection groups were removed and the product  
cleaved from the resin by the Low-High TFMSA procedure. The  
raw product was purified by preparative HPLC (reversed phase  
C<sub>18</sub> eluting with a gradient of A:0.1% TFA in water (MilliQ)  
and B:0.1%TFA, 10% water, 90% acetonitrile). The purified  
chimeric PNA-kemptide was characterized by analytical HPLC and  
25     FAB-MS.

Example 2

Radio labelling of the kemptide motif (Leu-Arg-Arg-Ala-Ser-  
30     Leu-Gly) by protein kinase A

The PNA-kemptide chimera of the formula:-

35      H-Leu-Arg-Arg-Ala-Ser-Leu-Gly-TGTACGTCACAACTA-NH<sub>2</sub>  
was labelled with <sup>32</sup>P in a reaction mixture containing:

PNA-kemptide, 10 $\mu$ M.....	5 $\mu$ l
10 x Protein Kinase A buffer.....	5 $\mu$ l
$\gamma$ $^{32}$ P ATP (>5000 Ci/mmol; 50 $\mu$ Ci/ $\mu$ l)....	10 $\mu$ l
Protein kinase A (Boehringer; 5mU/ $\mu$ l) .	0.2 $\mu$ l
5 H <sub>2</sub> O.....	30 $\mu$ l

The reaction was incubated for 30 minutes at 30°C and then for 10 minutes at 65°C before being centrifuged for 30 seconds at 15000g. The supernatant was transferred to a new eppendorff tube. Water was added to 1ml and the labelled PNA-kemptide was separated from unincorporated  $\gamma$   $^{32}$ P ATP by anion exchange chromatography using a DEAE Sephadex A-50 anion exchange column.

The specific activity of the PNA-kemptide was estimated at  $1 \times 10^8$  cpm/ $\mu$ g PNA-kemptide.

Example 3. Detection of DNA sequence via protection by PNA.

The 15-mer PNA bearing the radio-labelled kemptide motif as made in Example 2 was incubated for 15 minutes separately with each of the following nucleic acids:

1. None (See Fig. 1 lane 1)
2. Complementary DNA 15-mer. (See Fig. 1 lanes 2 and 5)
- 25 3. DNA 10-mer complementary to part of the sequence of the PNA. (See Fig. 1 lane 3)
4. DNA 40-mer 5' CTAGAGGATCTAGTTGTGACGTACAGGATCTTTTCATAG-3 containing a 15 base sequence complementary to the PNA in the centre of the sequence. (See Fig. 1 lanes 4 and 6)

In each case, 4 $\mu$ l of a 10 $\mu$ M solution of the oligonucleotide was mixed with 5 $\mu$ l of the labelled PNA in a 10 $\mu$ l reaction volume containing 30mM NaAc (pH 5.0), 50mM NaCl, 10mM ZnCl<sub>2</sub> and 5% v/v glycerol. Incubation for 15 minutes took place 35 at room temperature.

In the case of one out of two reactions using the 40-mer DNA and in the case of one out of two reactions using the 15-mer DNA, 15 units of mung bean nuclease were added and the reactions were incubated at 37° for 5 minutes.

5 All reactions were terminated by adding 10 $\mu$ l of loading buffer (30% v/v glycerol, 0.25% w/v bromophenol blue, 0.25% xylene cyanol, 0.01% SDS) and the reaction mixtures were subjected to electrophoresis in a 10% polyacrylamide gel and subsequently to autoradiography.

10 The results are shown in Figure 1 lanes 1 to 6. In lane 1 the labelled PNA is run alone and nothing is seen because it is immobile on the gel. In lane 2 the labelled PNA is hybridised to the 15-mer DNA and a band is seen for the hybrid. In lane 3 the labelled PNA is hybridised to the 10-mer DNA and it can be seen that the band produced is clearly distinguishable from that in lane 2 in the distance run on account of the difference in the size of the two oligo-DNA's involved. In lane 4 the PNA is hybridized to the 40-mer DNA and once again the distance run by the band on the gel allows 15 one to discriminate between the  $^{15}\text{PNA}/^{10}\text{DNA}$  and  $^{15}\text{PNA}/^{15}\text{DNA}$  hybrids and the  $^{15}\text{PNA}/^{40}\text{DNA}$  hybrid.

20 In lane 5 one sees the result of nuclease digestion in the case of the  $^{15}\text{PNA}/^{15}\text{DNA}$  hybrid. The band is the same as that in lane 2 showing that the nuclease has been unable to 25 degrade the hybridised DNA. Lane 6 shows the result of the nuclease digestion of the  $^{15}\text{PNA}/^{40}\text{DNA}$  hybrid. Once again, the band is as in lane 2, showing that only that part of the DNA which is hybridised to the PNA is protected against the nuclease. The internal 15-mer sequence in the 40-mer DNA 30 which is complementary to the PNA has therefore been detected by the procedure described above with reference to lane 6.

Example 4. Protection of RNA against nuclease by PNA.

35 PNA T<sub>10</sub> bearing the kemptide motif was labelled as in Example 2. The labelled PNA was incubated with a mixture of

polyadenylated RNA's according to the following procedure. PNA ( $5\mu\text{l}$ ) was incubated with  $4\mu\text{l}$  of  $100\text{ng}/\mu\text{l}$  mixed *in vitro* RNA transcripts (Gibco BRL) in a  $10\mu\text{l}$  reaction volume containing other constituents and under time and temperature  
5 conditions as in Example 3. A similar incubated mixture was further incubated with mung bean nuclease as in Example 3 and the reaction mixtures were prepared for and subjected to electrophoresis and autoradiographed as in Example 3. The results are seen in lanes 7 and 8 of Figure 1.

10 Lane 7 shows an unresolved ladder of different sizes of PNA/RNA hybrids. In lane 8, the nuclease has resolved the ladder into a single band by trimming off all the RNA's to the  $\text{A}_{10}$  portion protected by the PNA.

15 Example 5. Detection of single base pair variation in 15-mer DNA

20 The 15-mer labelled PNA used in Example 3 was hybridized to two different DNA's under both of two different temperature conditions as follows.

25 The labelled PNA was hybridised to complementary 15-mer DNA at incubation temperatures of  $55^\circ$  and  $65^\circ\text{C}$  respectively and in each case the reaction mixture was treated at that temperature with mung bean nuclease. In other respects the conditions were as in Example 3. Two similar incubations took place using the single base mismatched 15-mer DNA 5'-TAGTTGCGACGTACA -3' at the same two temperatures with subsequent addition of mung bean nuclease. The resulting autoradiographs are seen in lanes 9 to 12 of Figure 1.

30 In lane 9 the result of the hybridization of the PNA to its fully complementary DNA at  $65^\circ\text{C}$  with nuclease degradation is a band corresponding to protection of the DNA by the PNA. Lane 10 shows the result of hybridisation at  $65^\circ\text{C}$  with the single base mismatch. The DNA is not protected and no hybrid  
35 is seen. Lane 11 shows the result of the experiment at  $55^\circ\text{C}$  with the fully complementary sequence and lane 12 shows the

result of the 55°C experiment with the single mismatch. In both these lanes the protected DNA is detected.

The melting point for the perfectly matched 15-mer hybrid is 69°C. Under the stringent hybridisation conditions imposed by operating close to this melting point, the PNA distinguishes between the perfectly matched DNA and the single base mismatch for which the melting point is 61°C. At 55°C however, hybrids are formed both with the perfect match and the single base mismatch as would be predicted from the melting points of the hybrids.

Whilst the invention has been described with reference to the preferred practice thereof as illustrated by the above examples, many variations and modifications are possible within the scope of the invention.

CLAIMS

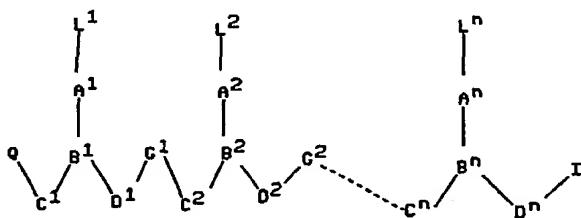
1. A method of protecting a selected region of a nucleic acid from attack by a reagent, which method comprises forming  
5 a complex between the nucleic acid and a nucleic acid analogue which hybridises thereto in a sequence selective manner and exposing said nucleic acid to a nucleic acid attacking reagent, wherein said complex is more stable against attack by said reagent than said starting nucleic acid.  
10
2. A method as claimed in claim 1, wherein said reagent is a nuclease.
3. A method as claimed in Claim 1 or Claim 2, wherein the  
15 nucleic acid analogue comprises a polymeric strand which includes a sequence of ligands bound to a backbone made up of linked backbone moieties, which analogue is capable of hybridisation to a nucleic acid of complementary sequence.  
20
4. A method as claimed in Claim 3, wherein said nucleic acid analogue backbone is a polyamide, polythioamide, polysulphinamide or polysulphonamide backbone.
5. A method as claimed in Claim 4, wherein said linked  
25 backbone moieties are peptide bonded amino acid moieties.
6. A method as claimed in any preceding claim, wherein the nucleic acid analogue is capable of hybridising to a nucleic acid of complementary sequence to form a hybrid which is more  
30 stable against denaturation by heat than a hybrid between the conventional deoxyribonucleotide corresponding in sequence to said analogue and said nucleic acid.
7. A method as claimed in any preceding claim, wherein said  
35 nucleic acid analogue is a peptide nucleic acid in which said backbone is a polyamide backbone, each said ligand being

bonded directly or indirectly to an aza nitrogen atom in said backbone, and said ligand bearing nitrogen atoms mainly being separated from one another in said backbone by from 4 to 8 intervening atoms.

5

8. A method as claimed in any preceding claim, wherein the nucleic acid analogue is capable of hybridising to a double stranded nucleic acid in which one strand has a sequence complementary to said analogue, in such a way as to displace 10 the other strand from said one strand.

9. A method as claimed in any preceding claim, wherein the nucleic acid analogue has the general formula 1:



Formula 1

wherein:

15 n is at least 2,

each of L<sup>1</sup>-L<sup>n</sup> is independently selected from the group consisting of hydrogen, hydroxy, (C<sub>1</sub>-C<sub>4</sub>) alkanoyl, naturally occurring nucleobases, non-naturally occurring nucleobases, aromatic moieties, DNA intercalators, nucleobase-binding groups, heterocyclic moieties, and reporter ligands;

each of C<sup>1</sup>-C<sup>n</sup> is (CR<sup>6</sup>R<sup>7</sup>)<sub>y</sub> where R<sup>6</sup> is hydrogen and R<sup>7</sup> is selected from the group consisting of the side chains of naturally occurring alpha amino acids, or R<sup>6</sup> and R<sup>7</sup> are independently selected from the group consisting of hydrogen,

25 (C<sub>2</sub>-C<sub>6</sub>) alkyl, aryl, aralkyl, heteroaryl, hydroxy, (C<sub>1</sub>-C<sub>6</sub>) alkoxy, (C<sub>1</sub>-C<sub>6</sub>) alkylthio, NR<sup>3</sup>R<sup>4</sup> and SR<sup>5</sup>, where R<sup>3</sup> and R<sup>4</sup>

are as defined below, and R<sup>5</sup> is hydrogen, (C<sub>1</sub>-C<sub>6</sub>)alkyl, hydroxy, alkoxy, or alkylthio-substituted (C<sub>1</sub> to C<sub>6</sub>)alkyl or R<sup>6</sup> and R<sup>7</sup> taken together complete an alicyclic or heterocyclic system;

5 each of D<sup>1</sup>-D<sup>n</sup> is (CR<sup>6</sup>R<sup>7</sup>)<sub>z</sub> where R<sup>6</sup> and R<sup>7</sup> are as defined above;

each of y and z is zero or an integer from 1 to 10, the sum y + z being from 2 to 10;

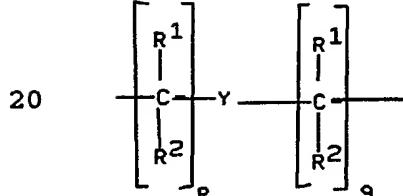
10 each of G<sup>1</sup>-G<sup>n-1</sup> is -NR<sup>3</sup>CO-, -NR<sup>3</sup>C<sup>S</sup>-, -NR<sup>3</sup>SO- or -NR<sup>3</sup>SO<sub>2</sub>-, in either orientation, where R<sup>3</sup> is as defined below;

each of A<sup>1</sup>-A<sup>n</sup> and B<sup>1</sup>-B<sup>n</sup> are selected such that:

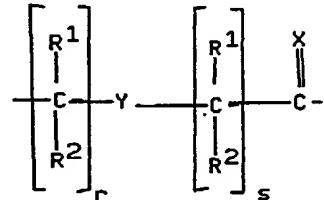
(a) A is a group of formula (IIa), (IIb), (IIc) or (IId), and B is N or R<sup>3</sup>N<sup>+</sup>; or

(b) A is a group of formula (IId) and B is CH;

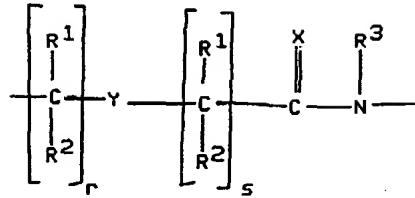
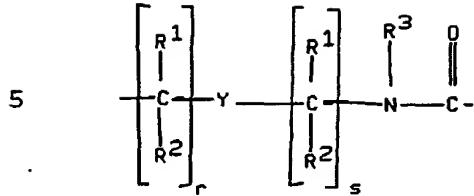
15



25      Formula IIIa



Formula IIb



10

Formula IIc

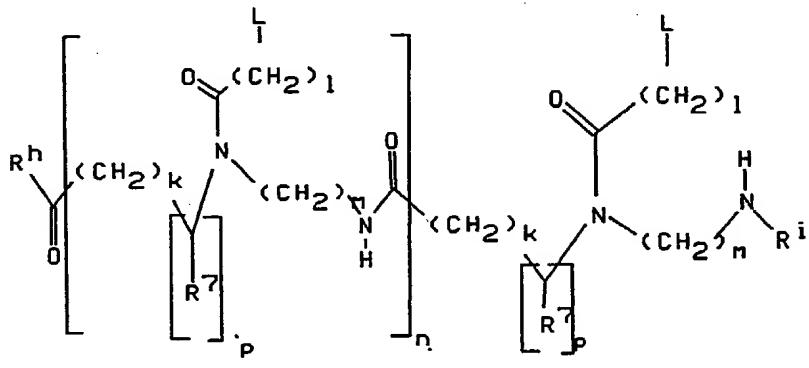
Formula IID

wherein:

- 15        X is O, S, Se, NR<sup>3</sup>, CH<sub>2</sub> or C(CH<sub>3</sub>)<sub>2</sub>;  
Y is a single bond, O, S or NR<sup>4</sup>;  
each of p and q is zero or an integer from 1 to 5, the  
sum p+q being not more than 10;  
each of r and s is zero or an integer from 1 to 5, the  
20 sum r+s being not more than 10;  
each R<sup>1</sup> and R<sup>2</sup> is independently selected from the group  
consisting of hydrogen, (C<sub>1</sub>-C<sub>4</sub>)alkyl which may be hydroxy- or  
alkoxy- or alkylthio-substituted, hydroxy, alkoxy, alkylthio,  
amino and halogen; and  
25        each R<sup>3</sup> and R<sup>4</sup> is independently selected from the group  
consisting of hydrogen, (C<sub>1</sub>-C<sub>4</sub>)alkyl, hydroxy- or alkoxy- or  
alkylthio-substituted (C<sub>1</sub>-C<sub>4</sub>)alkyl, hydroxy, alkoxy, alkylthio  
and amino;  
30        Q is -CO<sub>2</sub>H, -CONR'R'', -SO<sub>3</sub>H or -SO<sub>2</sub>NR'R'' or an  
activated derivative of -CO<sub>2</sub>H or -SO<sub>3</sub>H; and  
I is -NR'R'', where R' and R'' are independently selected  
from the group consisting of hydrogen, alkyl, amino protecting  
groups, reporter ligands, intercalators, chelators, peptides,  
proteins, carbohydrates, lipids, steroids, nucleosides,  
35        nucleotides, nucleotide diphosphates, nucleotide triphosphates,  
oligonucleotides, including both oligoribonucleotides

and oligodeoxyribo-nucleotides, oligonucleosides and soluble and non-soluble polymers.

10. A method as claimed in Claim 9, wherein said nucleic acid  
5 analogue comprises a compound of the general formula III, IV  
or V:

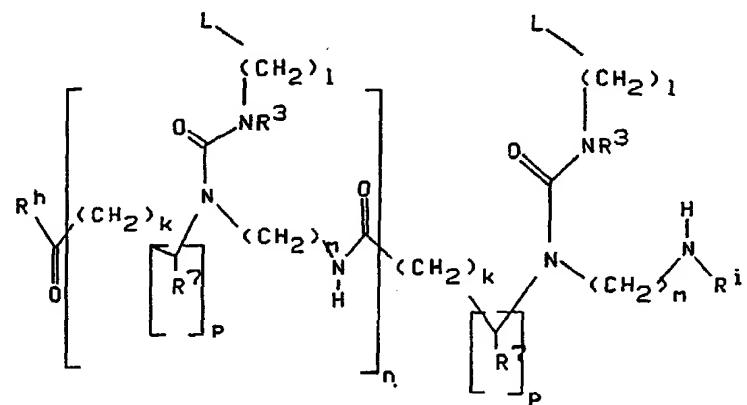


Formula III

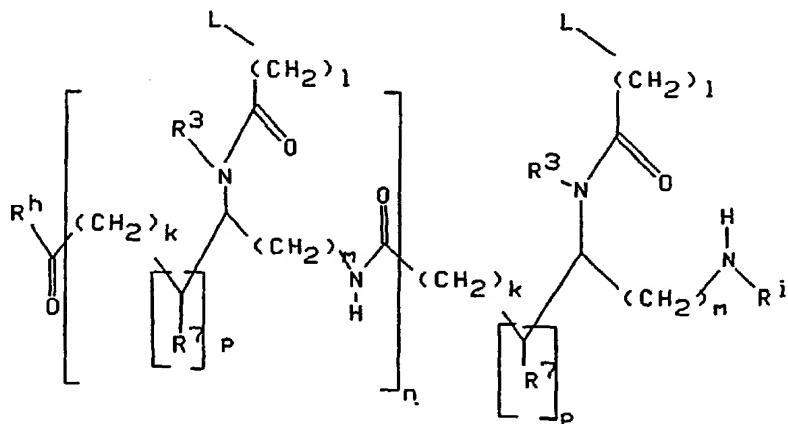
10

15

20



Formula IV



Formula V

wherein:

each  $\text{L}$  is independently selected from the group consisting of hydrogen, phenyl, heterocyclic moieties, naturally occurring nucleobases, and non-naturally occurring nucleobases;

5 each  $\text{R}^7$  is independently selected from the group consisting of hydrogen and the side chains of naturally occurring alpha amino acids;

n is an integer greater than 1,  
each k, l, and m is, independently, zero or an integer  
from 1 to 5;  
each p is zero or 1;  
5 R<sup>h</sup> is OH, NH<sub>2</sub> or -NHLysNH<sub>2</sub>; and  
R<sup>i</sup> H or COCH<sub>3</sub>.

11. A method of detecting the presence in a nucleic acid  
sample of a sequence, which method comprises exposing the  
10 sample of nucleic acid to a nucleic acid analogue capable of  
hybridising in sequence selective manner to said sequence if  
present in said nucleic acid sample under hybridising  
conditions so as to form a complex between said nucleic acid  
analogue and a region of said nucleic acid containing said  
15 sequence, exposing said nucleic acid to a reagent capable of  
degrading said nucleic acid under conditions such that the  
said sequence of said nucleic acid complex is protected from  
attack by said reagent whilst the remainder of said nucleic  
acid is degraded, and detecting the presence of said complex.

20 12. A method as claimed in claim 11, wherein the presence of  
said complex is detected by comparing the movement of said  
complex in electrophoresis with the movement of said nucleic  
acid analogue under similar conditions.

25 13. A method as claimed in claim 11 or claim 12, wherein the  
nucleic acid analogue bears a detectable label.

30 14. A method as claimed in any one of claims 11 to 13,  
wherein said label is a fluorescent label or a radio label.

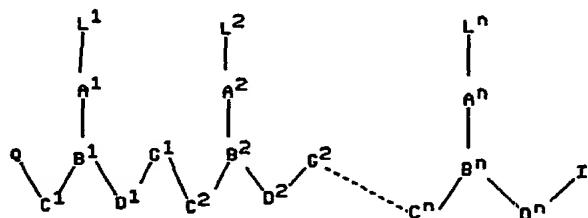
35 15. A method of carrying out a nucleic acid amplification  
comprising protecting a selected region of a nucleic acid  
within a nucleic acid sample by hybridising to said nucleic  
acid a nucleic acid analogue which hybridises thereto in a  
sequence specific manner, exposing said sample to a nucleic

acid attacking reagent which degrades the nucleic acid in the sample except for said selected region, and amplifying said selected region.

- 5    16. A method for working up the reaction product of a nucleic acid amplification procedure comprising protecting a selected region of the nucleic acid product of the amplification by forming a complex between the nucleic acid and a nucleic acid analogue which hybridises thereto in a sequence selective  
10 manner and exposing said nucleic acid to a nucleic acid attacking reagent, wherein said complex is more stable against attack by said reagent than said starting nucleic acid.
- 15    17. A method as claimed in Claim 16, wherein said selected region to be protected does not include at least one primer binding site necessary for the repetition of said amplification procedure.
- 20    18. A method as claimed in any one of Claims 11 to 17, wherein the nucleic acid analogue is capable of hybridising to a nucleic acid of complementary sequence to form a hybrid which is more stable against denaturation by heat than a hybrid between the conventional deoxyribo-nucleotide corresponding in sequence to said analogue and said nucleic  
25 acid.
- 30    19. A method as claimed in any one of Claims 11 to 18, wherein the nucleic acid analogue is a peptide nucleic acid in which said backbone is a polyamide backbone, each said ligand being bonded directly or indirectly to an aza nitrogen atom in said backbone, and said ligand bearing nitrogen atoms mainly being separated from one another in said backbone by from 4 to 8 intervening atoms.
- 35    20. A method as claimed in any one of Claims 11 to 19, wherein the nucleic acid analogue is capable of hybridising

to a double stranded nucleic acid in which one strand has a sequence complementary to said analogue, in such a way as to displace the other strand from said one strand.

- 5 21. A method as claimed in any one of Claims 11 to 20, wherein the nucleic acid analogue has the general formula 1:



Formula 1

wherein:

- n is at least 2,  
10 each of L<sup>1</sup>-L<sup>n</sup> is independently selected from the group consisting of hydrogen, hydroxy, (C<sub>1</sub>-C<sub>4</sub>) alkanoyl, naturally occurring nucleobases, non-naturally occurring nucleobases, aromatic moieties, DNA intercalators, nucleobase-binding groups, heterocyclic moieties and reporter;  
15 each of C<sup>1</sup>-C<sup>n</sup> is (CR<sup>6</sup>R<sup>7</sup>)<sub>y</sub> where R<sup>6</sup> is hydrogen and R<sup>7</sup> is selected from the group consisting of the side chains of naturally occurring alpha amino acids, or R<sup>6</sup> and R<sup>7</sup> are independently selected from the group consisting of hydrogen, (C<sub>2</sub>-C<sub>6</sub>) alkyl, aryl, aralkyl, heteroaryl, hydroxy, (C<sub>1</sub>-C<sub>6</sub>) alkoxy, (C<sub>1</sub>-C<sub>6</sub>) alkylthio, NR<sup>3</sup>R<sup>4</sup> and SR<sup>5</sup>, where R<sup>3</sup> and R<sup>4</sup> are as defined below, and R<sup>5</sup> is hydrogen, (C<sub>1</sub>-C<sub>6</sub>) alkyl, hydroxy, alkoxy, or alkylthio-substituted (C<sub>1</sub> to C<sub>6</sub>) alkyl or R<sup>6</sup> and R<sup>7</sup> taken together complete an alicyclic or heterocyclic system;  
20 each of D<sup>1</sup>-D<sup>n</sup> is (CR<sup>6</sup>R<sup>7</sup>)<sub>z</sub> where R<sup>6</sup> and R<sup>7</sup> are as defined above;

each of y and z is zero or an integer from 1 to 10, the sum  $y + z$  from 2 to 10;

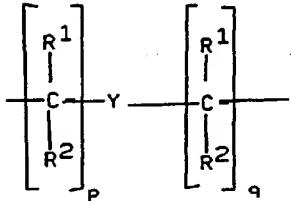
each of  $G^1-G^{n-1}$  is  $-NR^3CO-$ ,  $-NR^3C^5-$ ,  $-NR^3SO-$  or  $-NR^3SO_2-$ , in either orientation, where  $R^3$  is as defined below;

5 each of  $A^1-A^n$  and  $B^1-B^n$  are selected such that:

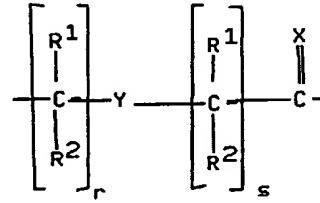
- (a) A is a group of formula (IIa), (IIb), (IIc) or (IId), and B is N or  $R^3N^+$ ; or  
(b) A is a group of formula (IId) and B is CH;

10

15



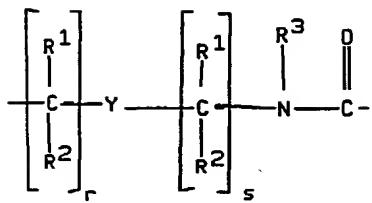
Formula IIa



Formula IIb

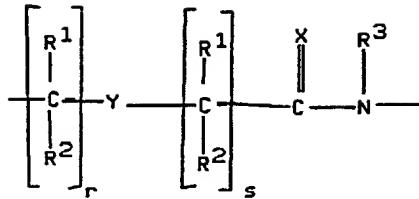
20

5



10

Formula IIc



Formula IID

15 wherein:

X is O, S, Se, NR<sup>3</sup>, CH<sub>2</sub> or C(CH<sub>3</sub>)<sub>2</sub>;

Y is a single bond, O, S or NR<sup>4</sup>;

each of p and q is zero or an integer from 1 to 5, the sum p+q being not more than 10;

20 each of r and s is zero or an integer from 1 to 5, the sum r+s being not more than 10;

each R<sup>1</sup> and R<sup>2</sup> is independently selected from the group consisting of hydrogen, (C<sub>1</sub>-C<sub>4</sub>)alkyl which may be hydroxy- or alkoxy- or alkylthio-substituted, hydroxy, alkoxy, alkylthio, amino and halogen; and

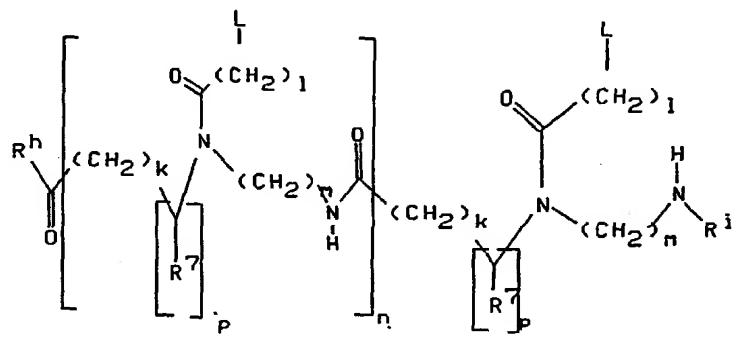
25 each R<sup>3</sup> and R<sup>4</sup> is independently selected from the group consisting of hydrogen, (C<sub>1</sub>-C<sub>4</sub>)alkyl, hydroxy- or alkoxy- or alkylthio-substituted (C<sub>1</sub>-C<sub>4</sub>)alkyl, hydroxy, alkoxy, alkylthio and amino;

30 Q is -CO<sub>2</sub>H, -CONR'R'', -SO<sub>3</sub>H or -SO<sub>2</sub>NR'R'' or an activated derivative of -CO<sub>2</sub>H or -SO<sub>3</sub>H; and

I is -NR'R'', where R' and R'' are independently selected from the group consisting of hydrogen, alkyl, amino protecting groups, reporter ligands, intercalators, chelators, peptides, 35 proteins, carbohydrates, lipids, steroids, nucleosides, nucleotides, nucleotide diphosphates, nucleotide triphos-

phates, oligonucleotides, including both oligoribonucleotides and oligodeoxyribonucleotides, oligonucleosides and soluble and non-soluble polymers.

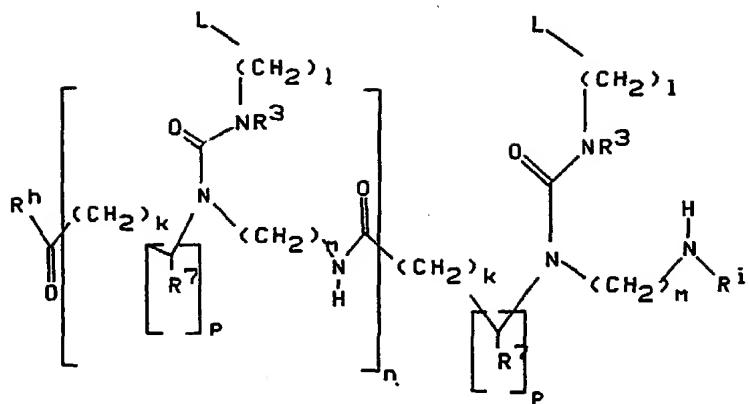
- 5 22. A method as claimed in Claim 21, wherein said nucleic acid analogue comprises a compound of the general formula III, IV or V:



Formula III

10

15



Formula IV

wherein:

each  $\text{L}$  is independently selected from the group consisting of hydrogen, phenyl, heterocyclic moieties,  
5 naturally occurring nucleobases, and non-naturally occurring nucleobases;

each  $\text{R}^7$  is independently selected from the group consisting of hydrogen and the side chains of naturally occurring alpha amino acids;

10  $n$  is an integer greater than 1,

each  $k$ ,  $l$ , and  $m$  is, independently, zero or an integer from 1 to 5;

each  $p$  is zero or 1;

$\text{R}^h$  is  $\text{OH}$ ,  $\text{NH}_2$  or  $-\text{NHLysNH}_2$ ; and

15  $\text{R}^i$  is  $\text{H}$  or  $\text{COCH}_3$ .

23. A method of detecting a nucleic acid sequence substantially as hereinbefore described in Example 3.

20 24. A method of protecting a nucleic acid from digestion by an attacking reagent substantially as hereinbefore described in Example 3 or Example 4 or Example 5.

25. A method of detecting a sequence variation substantially as hereinbefore described in Example 5.

Patents Act 1977  
 Examiner's report to the Comptroller under Section 17  
 (The Search report)

34

Application number  
 GB 9324955.5

**Relevant Technical Fields**

- (i) UK Cl (Ed.M) C3H (HB5, HB4B, HB7X) G1B (BAH, BAX)  
 (ii) Int Cl (Ed.5) C12N (15/11), C12Q (1/68)

**Databases (see below)**

- (i) UK Patent Office collections of GB, EP, WO and US patent specifications.  
 (ii) ONLINE DATABASES: WPI, DIALOG (BIOTECH)

Search Examiner  
 DR D ELSY

Date of completion of Search  
 8 MARCH 1994

Documents considered relevant following a search in respect of Claims :-  
 1-25

**Categories of documents**

- |    |   |    |   |
|----|---|----|---|
| X: | Document indicating lack of novelty or of inventive step.   | P: | Document published on or after the declared priority date but before the filing date of the present application.        |
| Y: | Document indicating lack of inventive step if combined with one or more other documents of the same category. | E: | Patent document published on or after, but with priority date earlier than, the filing date of the present application. |
| A: | Document indicating technological background and/or state of the art.   | &: | Member of the same patent family; corresponding document.   |

Category	Identity of document and relevant passages	Relevant to claim(s)
Y	WO 92/20703 A1 (BUCHARDT) see Claims	1-25
X,Y	HEALTH PERIODICALS ABSTRACT NO: 11677016, & SCIENCE, Vol 254, (6 December 1991), page 1497-1500 Nielsen et al	X: 1-25 Y: 1-25

Databases: The UK Patent Office database comprises classified collections of GB, EP, WO and US patent specifications as outlined periodically in the Official Journal (Patents). The on-line databases considered for search are also listed periodically in the Official Journal (Patents).